

This Page Is Inserted by IFW Operations  
and is not a part of the Official Record

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning documents *will not* correct images,  
please do not report the images to the  
Image Problem Mailbox.**

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2003, 15:04:29 ; Search time 17.2201 Seconds  
(without alignments)  
2160.495 Million cell updates/sec

Title: US-09-807-933b-9  
Perfect score: 2106  
Sequence: 1 MKFTVAITSAVALSSA.....TFKEVTCPAELTTRSGCERK 387  
Scoring table: BLOSUM62  
Gapop: 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues  
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_73.\*  
1: PIR1.\*  
2: PIR2.\*  
3: PIR3.\*  
4: PIR4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	758.5	36.0	229	JC7308	cellulase (EC 3.2.1.4)
2	592	28.1	511	S10527	endoglucanase B pr
3	525.5	25.0	393	S59499	cellulase eg11 - s
4	226.5	10.8	471	A26160	cellulose 1,4-beta
5	226.5	10.8	471	A38979	cellulose 1,4-beta
6	190	9.0	418	S28372	cellulase (EC 3.2.1.4)
7	186	8.8	410	S68153	cellulase (EC 3.2.1.4)
8	182	8.6	2395	S50820	surface protein ty
9	177	8.4	475	S49886	probable membrane
10	174.5	8.3	378	S61992	SLG1 protein - yea
11	174.5	8.3	533	T33912	hypothetical prote
12	172	8.2	438	S70602	cellulose 1,4-beta
13	171.5	8.1	1428	T08852	lustrin A - Califo
14	171	8.1	1777	T34369	hypothetical prote
15	168.5	8.0	402	T36185	hypothetical prote
16	168.5	8.0	4776	E95206	cell wall surface
17	166.5	7.9	481	A35628	loricrin - mouse
18	166.5	7.9	503	S63257	probable membrane
19	166	7.9	605	T33913	hypothetical prote
20	164	7.8	937	S58135	hypally regulated
21	163	7.7	600	S07638	spore coat protein
22	163	7.7	1275	T33369	hypothetical prote
23	163	7.7	1797	T21889	hypothetical prote
24	163	7.7	1805	T21888	hypothetical prote
25	161.5	7.7	354	T46740	microfilarial shea
26	161.5	7.7	662	A45155	mucin FIM-C.1 - Af
27	160	7.6	2543	T31687	surface antigen - P
28	159.5	7.6	2718	A23475	G surface protein
29	158.5	7.5	537	B33485	spore coat protein

30	158	7.5	786	1	A47547	serine proteinase
31	158	7.5	1022	2	T34433	hypothetical prote
32	158	7.5	1367	1	S48478	glucan 1,4-alpha-g
33	158	7.5	2232	2	T34434	hypothetical prote
34	156.5	7.4	2764	2	S09118	G surface protein
35	155.5	7.4	281	2	T49537	hypothetical prote
36	153	7.3	242	2	S60143	cellulase (EC 3.2.1.4)
37	152	7.2	463	2	T38444	hypothetical prote
38	151.5	7.2	962	2	S03818	carboxymethylcellu
39	151	7.2	738	2	S40992	hypothetical prote
40	151	7.2	739	2	B88553	protein K04H4.2b I
41	151	7.2	786	2	T16509	hypothetical prote
42	151	7.2	2639	2	T31328	fibroin - Chinese
43	150	7.1	13288	2	T03099	mucin, submaxillar
44	149.5	7.1	556	2	S51892	probable membrane
45	149.5	7.1	888	2	T46726	secreted acid phos

ALIGNMENTS

RESULT 1

JC7308  
Cellulase (EC 3.2.1.4) - Scopulariopsis brevicaulis  
N:Alternate names: endoglucanase I  
C:Species: Scopulariopsis brevicaulis  
C:Date: 18-Aug-2000 #sequence\_revision 18-Aug-2000 #text\_change 18-Aug-2000  
C:Accession: JC7308; PC7087  
R:Nakatani, F.; Kawaguchi, T.; Takada, G.; Sumitani, J.; Moriyama, Y.; Arai, M.  
Biosci. Biotechnol. Biochem. 64, 1238-1246, 2000  
A:Title: Cloning and sequencing of an endoglucanase gene from Scopulariopsis brevicaulis  
A:Reference number: JC7308  
A:Accession: JC7308  
A:Molecule type: DNA  
A:Residues: 1-229 <NAK>  
A:Experimental source: strain TOF-1212  
A:Accession: PC7087  
A:Molecule type: protein  
A:Residues: 21-37;149-164 <NA2>  
C:Genetics:  
A:Gene: eg1  
A:Introns: 147/3  
C:Keywords: glycosidase; hydrolase

Query Match	36.0%	Score 758.5;	DB 2;	Length 229;
Best Local Similarity	62.8%;	Pred. No. 7.1e-39;		
Matches 130;	Conservative 30;	Mismatches 44;	Indels 3;	Gaps 3;
Qy	183	SGSGSTRTRYDCCCKASCSWPGKASVT-GPVDTCASNGISLLDAN-AOSGCN-GGNGFMCN	239	
Db	22	SGTGTTRYDCCCKSPCSWPKAPLSQGPMTCTDINDNPLDDGGLTESGCEPGGAYMCS	81	
Qy	240	NNQWAVNDELAYGFAAASIAGSNEAGWCCGYELTFTSGAASCKMNVQVNTTGGDLGS	299	
Db	82	SHSPWAVDDELLAYGAAAVNIGGQTESDMCCACYELEFTTGAIVSKMIVQATNTGGDLGN	141	
Qy	300	NHFDLPQGGGVGIFNGCAAGWPNQWAGRYGVSVSDCASLPSALQAGCKWRNWF	359	
Db	142	NHFDIAPQGGGVGIFNGCTDQWSPNGWDRYGVVHTRADCDSPFPAKAGCEWRDWF	201	
Qy	360	KNSDNPMTTFKVTCPAELTTRSGCER	386	
Db	202	GGTDNDPVSFREVCECPAELVQKSQCR	228	

RESULT 2

S10527  
endoglucanase B precursor - Pseudomonas fluorescens  
C:Species: Pseudomonas fluorescens  
C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 08-Oct-1999  
C:Accession: S10527  
R:Gilbert, H.J.; Hall, J.; Hazlewood, G.P.; Ferreira, L.M.A.  
Mol. Microbiol. 4, 759-767, 1990

	Query Match	25.0%	Score 525.5	DB 2	Length 393
	Best Local Similarity	48.6%	Pred. No. 1.1e-24		
	Matches 107	Conservative 28	Mismatches 60	Indels 25	Gaps 10
Qy	185	SGSTTRVYWDCCKASCSPFGKASVTPVDVTCASNGISILDA----	NAOSGCGNGGNGFNCNN	240	
Db	26	AGMATRVYWDCCLASASMEGKAPVAPVDACADAGVTLLDSKKDPSGGSGGCGNGKFKWCSC	85		

Query Match	10.9%;	Score 226.5;	DB 1;	Length 471;			
Best Local Similarity	25.1%;	Pred. No. 1.1e-06;					
Matches 125;	Conservative 37;	Mismatches 145;	Indels 191;	Gaps 30;			
Qy	7	ITSIAVALALSSSA---	EAASCSVYGQCGIGWTGPTCCDAGSTCKAQRDNKYYSCIP	63			
Db	6	LTTLATLTLAAVPLEERQACSVMGCGQWWSGPTCCASGSTCVSND--	YYSQCLP	63			
Qy	64	KPGSSSSSSCSVYGQCGIGWSGPTCCES-----	GSTCVAQEGNKY	106			
Db	64	GAASSSSSTRAASTTSRV-----	SPTTSRSSATPPGSTTTTRVPPVGSGTATYSGNPF	117			
Qy	107	YSQCLPCSHNNAGNASTTKTS-TKTSTTTAKATATVTKTKTKTKTKTTTSTA	AAAA	165			
Db	118	VG-VTPMANAYAYASEVSSLAIPSLUTGAMATAAAVAKVPSFMWLDLTKTPLMEQTLADI	176				
Qy	166	STSTSSAGYKVISGGKSGSGSTTRY-----	WDCC-----KASCSPGKASVTGTPVDTC	215			
Db	177	RTANKGGNY-----	AGQFWYDLPDRDCAALASNGEYSIADGGVAKYKNYIDTIR	227			
Qy	216	SGNISLLD-----	ANAQSGCGNGGFMCMN-----	NQP--243			
Db	228	QIVVEYSDIRTLVIEPDSLNLVNLGTPKCANAOQAY-----	LECINYAVTOLNLPNV	282			
Qy	244	-----	WAVN-DELAYGFA-----	AASAGSNEAGWCCGCVEL	274		
Db	283	AMYLDAHAGWLGWPANQDPAAQLFANVYKNASSPRALRGLATNVANYN--	GW-----	NI	335		
Qy	275	T-----	FTSG-AAAGKKMVMVQVTTGTGDLGSNH-----	FDL-----	OMPGGGVGI	FNCG	317
Db	336	TSPSYTQGNVAVNEKLYIHAI---	GPLLANHGWSNAAFFITDGRSGKQPTG-----	---	---	---	384

```
Qy 318 AAQMG-----APNDG-----W-----GARYGVSSV-----SDCASLPS 346
|||
Db 385 QQQGDMCNVIGTGFGRPSANTGSLDLSFVWVKPGGECGTGSDSSAPRFDHCA-LPD 443
|||
Qy 347 AL-----QAGCKWRFNWFK 360
|||
Db 444 ALQAPAQAGA-----WFO 456
|||

RESULT 5
A38979
cellulase 1,4-beta-cellobiosidase (EC 3.2.1.91) II - fungus (Trichoderma viride)
C:Species: Trichoderma viride
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: A38979
R:Wang, J.R.; Zhang, M.F.; Huang, T.
Acta Genet. Sin. 22, 74-80, 1995
A:Title: The primary structure of cellobiohydrazase gene (CBH II) from Trichoderma viride.
A:Reference number: A38979
A:Accession: A38979
A:Molecule type: DNA
A:Residues: 1-471 <WAV>
C:Genetics:
A:Gene: cbhII
A:Introns: 31/2; 160/2; 243/1
C:Superfamily: cellulase 1,4-beta-cellobiosidase II; fungal cellulose-binding domain hom
C:Keywords: glycosidase; hydrolase; polysaccharide degradation
F:31-62/Domain: fungal cellulose-binding domain homology <PCB>

Query Match 10.8%; Score 226.5; DB 1; Length 471;
Best Local Similarity 25.1%; Pred. No. 1.1e-06;
Matches 125; Conservative 37; Mismatches 145; Indels 191; Gaps 30;

Qy 7 ITSIAVALALSSA---BAACSSVYGCGGIGWTGPTCCDAGSTCKAQKKNKYQCIP 63
|||
Db 6 LTTLATLATLAASVPLEERQACSSVWGCGGQNSGPTCCASGSTCVYSND--YYSQCLP 63
|||
Qy 64 KPGKSSSSSSSVSYQCIGWSGPTCCES-----GSTCVAQEGNKY 106
|||
Db 64 GAASSSSSTRAASTTSRV-----SPTTSRSSSATPPPGSTTRVPVVGSGTATYVGNPF 117
|||
Qy 107 YSQLPGSHSNAGNASSTKRTS-TKTSTTTAKATATVTTTKTTKTTTKTTTSTA 165
|||
Db 118 VG-VTPMANAYVASEVSLAIPSLTGAMATAAAAVKVPFWMFLDTLTKPLMEQTLADI 176
|||
Qy 166 STSTSSSAGYKVISGKSGSGSTTRY-----WDCC-----KASCSWPCKASVTGPVDTCA 215
|||
Db 177 RTANKNGNY-----AGQFVVYDLPDRDCAALASNGEYSIADGGVAKYKNYIDTIR 227
|||
Qy 216 SNGISLLD-----ANAQSCGNGGFMCKNN-----NQP-- 243
|||
Db 228 QIVVEYSIDRTLLVTEPDSLANLVNLGTPKCANQAQAY-----LECINYAVTQLNPNV 282
|||
Qy 244 -----WAVN--DELA YGFA-----AASAGSNEAGWCCGCGYEL 274
|||
Db 263 AMYLDAGHAGWLGWPAQDPAQLFANYVKNASSPRALRGLATNVANYN--GW-----NI 335
|||
Qy 275 T-----FTSG-AASGKKVMVQVNTNGDILGSHN-----FDL-----QMPGGGVGIFNGC 317
|||
Db 336 TSPSPSYTCGNVAYNEKLYIHAI--GPLLNHGNSNAFFITDQGRSGKQPTG----- 384
|||
Qy 318 AAQMG-----APNDG-----W-----GARYGVSSV-----SDCASLPS 346
|||
Db 385 QQQGDMCNVIGTGFGRPSANTGSLDLSFVWVKPGGECGTGSDSSAPRFDHCA-LPD 443
|||
Qy 347 AL-----QAGCKWRFNWFK 360
|||
Db 444 ALQAPAQAGA-----WFO 456
|||

RESULT 6
S28372
```

```
cellulase (EC 3.2.1.4) III precursor - fungus (Trichoderma reesei)
N:Alternate names: endo-1,4-beta-glucanase; endoglucanase III
C:Species: Trichoderma reesei
C:Date: 11-Sep-1998 #sequence_revision 11-Sep-1998 #text_change 18-Jun-1999
C:Accession: S28372; S78516; S78517; S02626
R:Salohelmo, M.; Lehtovaara, P.; Penttilae, M.; Teeri, T.T.; Stahlberg, J.; Johansson, G
Gene 63, 11-21, 1988
A:Title: EGTIII, a new endoglucanase from Trichoderma reesei: the characterization of bot
A:Reference number: S28372; MUID:98255850; PMID:3384334
A:Accession: S28372
A:Molecule type: DNA
A:Residues: 1-418 <SAL1>
A:Cross-references: EMBL:M19373; NID:g170548; PIDN:AAA34213.1; PID:g170549
A:Accession: S78516
A:Molecule type: mRNA
A:Residues: 1-418 <SAL2>
A:Cross-references: GB:M19373; NID:g170548; PIDN:AAA34213.1; PID:g170549
A:Accession: S78517
A:Molecule type: protein
A:Residues: 22-23,'X',25-46,'XX',49-58 <SAL3>
R:Stahlberg, J.; Johansson, G.; Pettersson, G.
Eur. J. Biochem. 173, 179-183, 1988
A:Title: A binding-site-deficient, catalytically active, core protein of endoglucanase I
A:Reference number: S02625; MUID:88185316; PMID:3356188
A:Accession: S02626
A:Molecule type: protein
A:Residues: 'XXX',87-88,'X',92-93,'Y',95-99 <STA>
A:Experimental source: culture filtrate
C:Genetics:
A:Gene: egI3
A:Introns: 110/2
C:Function:
A:Description: hydrolysis of 1,4-beta-D-glucosidic linkages in beta-D-glucans such as ce
A:Pathway: cellulose degradation
C:Superfamily: Trichoderma cellulase III; fungal cellulose-binding domain homology
C:Keywords: blocked amino end; glycoprotein; glycosidase; hydrolase; polysaccharide degr
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-418/Product: cellulase #status experimental <MAT1>
F:26-57/Domain: fungal cellulose-binding domain homology <PCB>
F:83-418/Product: cellulase core #status experimental <MAT2>
F:22/Modified site: carboxylic acid (Gln) (in mature form) #status experimental
F:124/Binding site: carboxylate (Asn) (covalent) #status predicted

Query Match 9.0%; Score 190; DB 1; Length 418;
Best Local Similarity 24.3%; Pred. No. 0.00015;
Matches 83; Conservative 41; Mismatches 119; Indels 98; Gaps 15;

Qy 76 SVYSQCGIGWSGPTCCESGTCVAQEGNKYYSQCLPGSHSNAGNASSTKTKTSTTT 135
|||
Db 24 TVMQCGGIGWSGPTCCASGACSTL--NPYYAQICPA----- 60
|||
Qy 136 TAKATATVTTTKTTKTTTKTTTAAASTSTSSSAGYKVISGKSGSGSTTRYWDCC 195
|||
Db 61 -----TTTTSTRPPSGPTTTTATSTSSSTSSGVRFAGVNIAGDFGCTTD--GTC 113
|||
Qy 196 KASCSWPCKASVTGPVDTCSN-----GISLIDANAQSCGNGGFMCKNNQPMWVDELA 251
|||
Db 114 VTSKYVPLKNFTG-----SNYPDGIQGMQHFNE--DGMTIFRLPVGHQYLVNNILG 165
|||
Qy 252 YGFAAASAGSNEAGWCCGCGYELFTTSCAASGKKVMVQVNTN-----TGCDLGSNHFDLQMP 307
|||
Db 166 GNLDSSTISKYDQL--VQGCLSL-----GAYCIVDIHNYARWNGGIQ----- 207
|||
Qy 308 GGG-----VGIFNGCAAQWGA-----PND-----GWARYGV-----SSVS 339
|||
Db 208 -GGPTNAQFTSLWSQLASKYASQSRVMFGINMNEPHDVIINTWAAATVQEWVTAIRNAGATS 266
|||
Qy 340 DCASLPSALQAGCKWRFNWFKNSDNTWTTFKEVTCPAELTT 380
|||
Db 267 QFISLP-----GNDWQSAQAFISDGSAAALSQVTNPDGSTT 302
|||

RESULT 7
```

Matches	107;	Conservative	43;	Mismatches	192;	Indels	130;	Gaps	24;																											
QY	2	KFTVAITSI	AV	AL	ALLSS	GA	EA	AA	SC	SV	Y	Q	Q	CGG	---	IG	WT	GP	TC	---	CD	AG	ST	C	49											
DB	1857	KSTCSA	VT	IT	DT	VA	LL	AL	NG	TV	CA	MD	SA	Q	K	CR	DC	DF	GS	TT	HA	AC	Q	T	OR	AG	CT	AG	AG	1916						
QY	50	KAOK	---	DN	KY	SO	CI	P	K	P	K	G	---	SS	SS	SS	SS	SS	SS	SS	SS	SS	SS	SS	SS	SS	SS	SS	SS	93						
DB	1917	KCARV	Q	NC	EG	TT	IR	SA	CT	EG	NP	CL	IN	NY	V	NS	D	G	K	G	K	A	F	R	T	S	C	K	S	1976						
QY	94	---	SG	ST	CV	---	AO	EG	NK	---	Y	S	O	C	L	P	E	S	H	N	N	A	G	N	A	S	S	T	K	T	135					
DB	1977	ISNO	CT	NG	S	NC	IG	IT	LC	SE	T	T	D	G	C	V	G	D	GA	CI	---	OS	V	P	A	L	N	S	S	D	2032					
QY	136	TAKAT	A	T	A	T	V	T	K	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	184						
DB	2033	---	AD	A	P	Y	T	H	S	D	Q	T	A	S	K	K	T	T	NG	T	T	G	C	I	A	L	G	A	C	S	2089					
QY	185	SG	ST	R	Y	W	D	C	K	A	S	C	S	W	P	K	A	S	V	T	G	---	P	V	D	T	C	A	S	N	G	2306				
DB	2090	TG	I	C	T	---	WD	T	A	S	S	C	R	D	Q	S	C	A	D	L	T	G	T	T	H	A	T	C	S	S	317					
QY	231	N	---	GG	N	F	M	C	---	NN	Q	P	H	---	V	N	D	E	L	A	Y	G	F	A	A	S	A	G	S	N	E	269				
DB	2148	TTA	V	G	S	D	---	I	C	Y	W	E	L	A	S	T	NN	N	T	A	K	R	L	I	A	C	I	O	NG	T	S	T	2206			
QY	270	-G	C	V	E	L	T	---	F	T	S	G	A	S	G	K	M	V	V	T	T	G	D	L	S	N	H	F	D	L	Q	M	P	317		
DB	2207	A	N	C	S	T	Y	T	T	K	T	A	C	N	S	G	L	D	I	C	V	F	T	O	S	T	A	T	A	G	A	A	G	T	2257	
QY	318	-A	A	Q	M	A	P	N	D	G	W	A	R	Y	G	---	V	S	S	V	---	S	L	P	S	A	L	O	A	C	K	W	R	F	362	
DB	2258	V	A	C	O	A	K	D	R	C	S	W	T	A	A	S	G	T	G	T	T	A	V	A	S	K	A	T	H	T	C	A	T	N	G	2309

RESULT 9

S49886

probable membrane protein YIL123w - yeast (Saccharomyces cerevisiae)

N:Alternate names: hypothetical protein YI8277.06

C:Species: Saccharomyces cerevisiae

C:Date: 13-Jan-1995 #sequence\_revision 20-Feb-1995 #text\_change 11-Jan-2000

C:Accession: S49886

R:Hamlyn, N.; Churcher, C.

submitted to the EMBL Data Library, November 1994

A:Reference number: S49881

A:Accession: S49886

A:Molecule type: DNA

A:Residues: 1-475 <HAM>

A:Cross-references: GB:247047; EMBL:246833; NID:g603997; PID:g763223; MIPS:YIL123w

C:Genetics:

A:Gene: SGD:SIM1

A:Cross-references: SGD:S0001385; MIPS:YIL123w

A:Map position: 9L

C:Superfamily: Saccharomyces NCA3 protein

C:Keywords: transmembrane protein

F:82-98/Domain: transmembrane #status predicted <TMM>

Query Match

Best Local Similarity

Matches

8.4%; Score 177; DB 2; Length 475;

24.6%; Pred. No. 0.001;

51; Mismatches 139; Indels 110; Gaps 18;

QY	4	TV	A	I	T	S	I	A	V	A	L	S	S	A	E	A	A	S	C	S	S	V	Y	Q	Q	G	G	I	G	T	G	T	C	C	D	A	G	S	T	K	A	O	K	N	K	Y	---	Y	S	59
DB	67	T	S	A	I	T	T	S	V	---	L	A	P	T	S	S	A	A	A	G	I	A	S	I	---	---	A	V	S	S	A	A	L	A	K	N	E	K	I	S										



[illegible]

db 1334 CRTWL 1338

RESULT 14

T34369

hypothetical protein T19D12.1 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 29-Oct-1999

C;Accession: T34369

R;Ravello, A.

submitted to the EMBL Data Library, November 1995

A;Description: The sequence of C. elegans cosmid T19D12.

A;Reference number: Z21513

A;Accession: T34369

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-1777 <FAV>

A;Cross-references: EMBL:U41263; PIDN:AAC24428.1; GSPDB:GN00020; CESP:T19D12.1

A;Experimental source: strain Bristol N2; clone T19D12

C;Genetics:

A;Map position: 2

A;Introns: 36/1; 134/2; 180/1; 622/3; 691/2; 754/1; 1111/2; 1174/1; 1271/3; 1334/2

Query Match 8.1%; Score 171; DB 2; Length 1777;

Best Local Similarity 27.9%; Pred. No. 0.0073;

Matches 72; Conservative 37; Mismatches 101; Indels 48; Gaps 12

QY 17 SSSAEAAACSSVYGCGGIGHTGPTCCDAGSTCKAOKDNKYVQCIPKPGKSSSSSSSSSS 76

DB 1487 SSSIQPIASSSTMGTAGSSSPQPTVS-----STVSPSTGATSSGSSSTT 153

QY 77 VYSCQCGICWGSPTCCSGSTCVAGENKYYVQCIPGSHNNAGNASTKTKSTTKSTTT 136

DB 1532 V-----GSSTVGATQTSVSSVTPVNTGTTGVTNPSTSSSTSG-SSSTQSPSSTAANT 158

QY 137 AKATA--TVTTKTVKTKTKTKTKTSTAATASTSSAGYKVIKSGKSGSGSTTRYWDC 194

DB 1586 GSSTSGPTVAT-TQGSSTQTNSNTGSGTTVATTVQSS-----TGNSSG--TTVNFQSS 163

QY 195 CKASCSPGKASV---TGPVDTCAS--NG--ISLIDANAQCGNGGFMCNNQPAVN 247

DB 1639 SVASTVASSTANESSTTPIPCASPYSGKIATVFELNANSATVNV-FVSNN----- 169

QY 248 DELAYGFAAASIAGSNEA 265

DB 1691 ---LYNSANYDFTNSTEA 1705

RESULT 15

E86185

hypothetical protein [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001

C;Accession: E86185

R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, C.

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.;

ansen, N.P.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, I.

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, Z.A.; Luros, J.S.; Maiti, R.;

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A;Reference number: AB6141; MUID:21016719; PMID:11130712

A;Accession: E86185

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-402 <STO>

A;Cross-references: GB:AE005172; NID:Q2388584; PIDN:AAB71465.1; GSPDB:GN00141

C;Genetics:



A;Map position: 1

```
Query Match      8.0%; Score 168.5; DB 2; Length 402;
Best Local Similarity 27.8%; Pred. No. 0.0029;
Matches 84; Conservative 27; Mismatches 132; Indels 59; Gaps 12;

Qy 41 TCCDAGSTCKAOKNKYISQCIPKPGSSSSSSSVYQCGGIGWSG-PTCCSGSTCV 99
Db 9 TASAAGTLTRASTYPNTYNSNPADASSMPSTWTTVPQTSSSSTSTTTATSSSGTT 68

Qy 100 AOEKNKYISQCLPCS--HSNNAGNASSTKKT-----STKST-----TTAK 138
Db 69 AESSSTKSAITMSGSTTHTTSSATASSTASTSTSYSTSYSTSKTTTGTGTISTTAS 128

Qy 139 ATATVTTKTVTK-----TTTKTTKT-----STTAAASTSTSSAGYKVIS---GGK 182
Db 129 AAPTSTASTSTSYSTSYSTSKTTTGTGTIGTTASAAPTSTSTANSASSTNPS 188

Qy 183 SGSGSTTRYNDCKKASCSWP-GKASVTGPDVTC--SNG-----ISLLDANAQGCNGN 234
Db 189 SGSKPTAMTGTTANTSPSAPTSPSTTNSSSTAAYTSSGSKPTTTRTTANTSSASTSS 248

Qy 235 GFMCNNOPWAVNDELAYGFAAASIAGSNE---AGWCCGCYELTFTGAASCKKVVQVT 291
Db 249 ASPTNSSTSTPTN-----SSAGSKPTTMTGTTTNTSTSTTTSSASTKSSSSAT 298

Qy 292 NT 293
Db 299 NS 300
```

Search completed: June 18, 2003, 17:16:19  
Job time : 18.2201 secs

**THIS PAGE BLANK (USPTO)**